Unveiling Insights: A Deep Dive into H1B Visa Approvals Over 8 Years

As H1B visa approvals remain a focal point in contemporary discussions, particularly amidst the increasing foreign population in the U.S. (myself included), I embarked on my first data analytics project. My aim? To delve into the numbers and uncover insights into this significant aspect of immigration.

H1B visa is a non-immigrant visa that allows U.S. employers to temporarily employ foreign workers in specialty occupations. These occupations typically require specialized knowledge and a bachelor's degree or higher. H1B visa approvals are essential in maintaining a balance between employment of non-immigrant workers and US workers. It fosters diversity, inclusion, and improves economic growths as well.

Purpose of this project:

I was drawn to the complexity of this concept. Once I started digging into it, I understood that I had to work with a lot of data, which serves the purpose of learning and finding my way to arrive at meaningful insights.

By analysing the data available in USCIS and the US department of labor, I was able to arrive at insights like percentage of H1B approvals and denials, identified top employers, job titles, and industries spanning over an 8-year period.

Data Collection and Preparation:

- I began my data journey by cleaning the raw data using Excel. This involved removing irrelevant information, converting general values into numerical formats to prevent conflicts during SQL import, and addressing empty cells that could disrupt analysis.
- As a Mac user, I utilized MySQL Workbench for SQL queries, although I encountered challenges when importing data from Excel due to the file's size. To overcome this, I opted for loading data via the terminal, a process detailed here.
- Once the data was imported, I applied various SQL techniques including aggregation, sorting, sub-queries, temporary tables, and joins to manipulate and analyse the dataset effectively.

Important SQL queries:

```
--To list the top 10 employers ordered by approval percentage

with total_initial_approval as
(select sum(initial_approval) as total from h1b_8years)
select employer_name, industry_code, sum(initial_approval) as approval_sum,
round((sum(initial_approval)/(select total from total_initial_approval)) * 100,2) as
approval_percent
from h1b_8years
group by employer_name, industry_code
order by 3 desc limit 10;

--To list the top 10 job titles that the top 10 employers hire for

create temporary table temp1
select employer_name, sum(initial_approval) as approval_sum
from h1b_8years
group by employer_name
order by 2 desc limit 10;

Select A.soc_title, count(A.soc_title) from job_title A join temp1 B
on A.employer_name = B.employer_name
group by A.soc_title
order by 2 desc limit 10;
```

Then, I exported the results from MySQL workbench and loaded them into Tableau for visualization. You can explore all the findings in my Tableau dashboard here.

Key findings:

According to the last 8 years(2018–2024) of data from USCIS, these are the top 10 employers in US who approve H1B visa for non-immigrants.

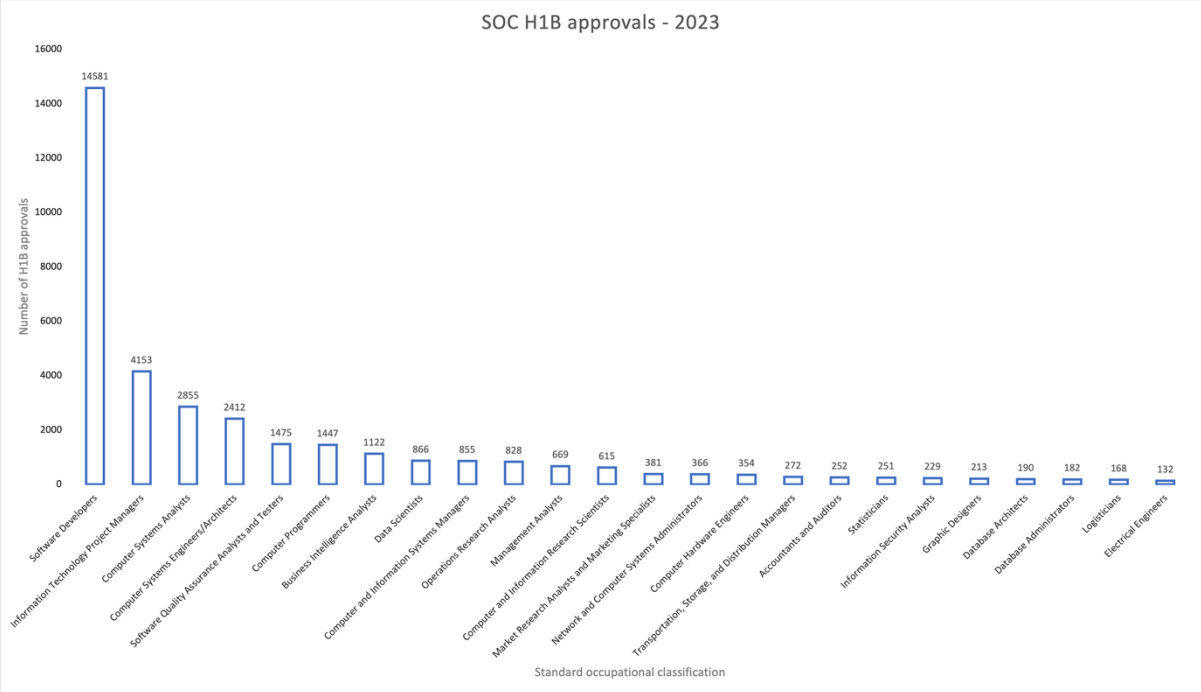
Approval by Industry codes

Industry Code	
11 - Agriculture, Forestry, Fishing and Hunting	0.06%
21 - Mining, Quarrying, and Oil and Gas Extraction	0.18%
22 - Utilities	0.20%
23 - Construction	0.76%
31-33 - Manufacturing	9.22%
42 - Wholesale Trade	0.83%
44-45 - Retail Trade	3.01%
48-49 - Transportation and Warehousing	0.44%
51 - Information	5.81%
52 - Finance and Insurance	5.38%
53 - Real Estate and Rental and Leasing	0.36%
54 - Professional, Scientific, and Technical Services	50.19%
55 - Management of Companies and Enterprises	0.43%
56 - Administrative and Support and Waste Manageme..	1.04%
61 - Educational Services	10.75%
62 - Health Care and Social Assistance	5.33%
71 - Arts, Entertainment, and Recreation	0.14%
72 - Accommodation and Food Services	0.18%
81 - Other Services (except Public Administration)	0.38%
92 - Public Administration	0.17%
other	5.13%

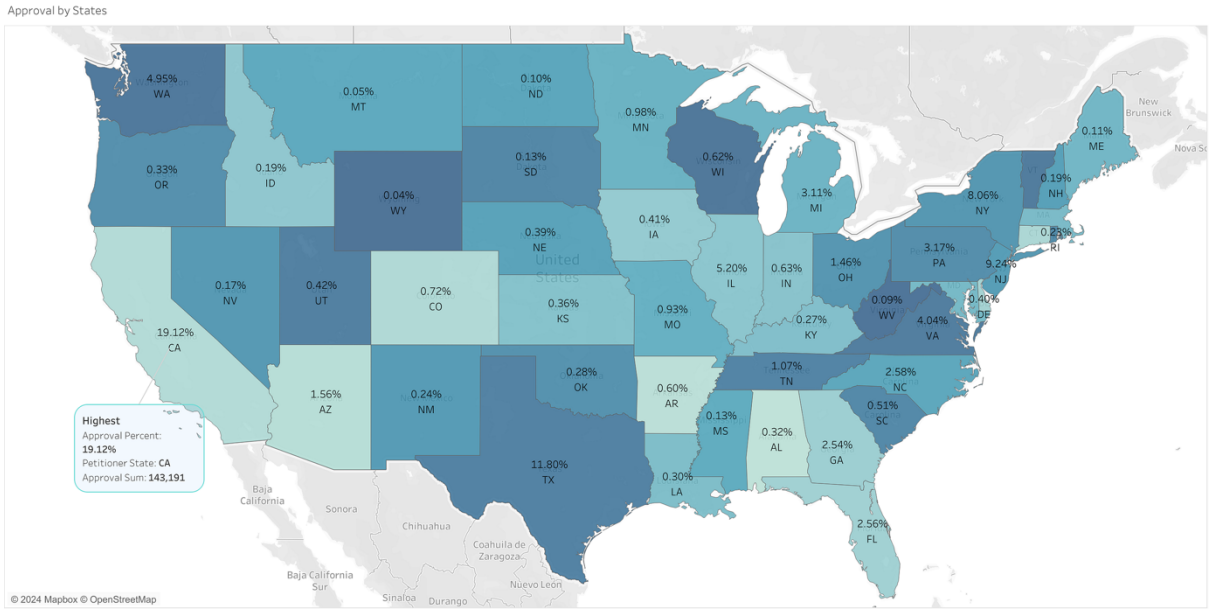
Top 10 employers

Employer Name	
INFOSYS LIMITED	1.80%
TATA CONSULTANCY SVC..	1.59%
GOOGLE LLC	1.27%
MICROSOFT CORPORATION	1.09%
AMAZON.COM SERVICES ..	0.92%
IBM CORPORATION	0.88%
WIPRO LIMITED	0.75%
APPLE INC	0.72%
COGNIZANT TECHNOLOG..	0.72%
CAPGEMINI AMERICA INC	0.71%

With **'Professional, scientific and technical services'** being the industry that subjects to more than 50 percent of the approvals, these are the top 25 job roles for which H1B visas are approved.



In the last 8 years, data from USCIS highlights computer science roles like **developers, analysts, and architects** as having higher H1B approval rates, emphasizing the industry's preference for specialised skills in this field.



California leads the list of states with the highest H1B visa approvals, followed by **Texas, New Jersey, New York, Illinois, Washington, and Massachusetts.** For more details, you can hover over the states in the dashboard here.

Using Tableau's forecasting feature, we observe a rise in approval percentages and a decline in denial percentages over time. Looks like a win-win situation for foreign workers—just kidding!

Wrapping up my first data analytics project on H1B visa approvals, I acknowledge the potential for further insights from the available data. Nonetheless, I'm pleased with the outcomes achieved and the valuable lessons learned along the way. I invite readers to share their suggestions, questions, and feedback, as continuous improvement is key to refining our understanding of complex datasets.

References:

Links to where you can download raw data,

- https://www.uscis.gov/tools/reports-and-studies/h-1b-employer-data-hub
- https://www.dol.gov/agencies/eta/foreign-labor/performance

You can also download the cleaned and processed data here.